Amendments to the Specification:

Please replace paragraph 8 beginning on page 4, with the following amended paragraph:

Another method of creating a ring tone is to translate recorded music into a sequence of tones. There are a number of problems that arise when attempting to translate recorded music into a ring tone sequence for an electronic device. The translation process generally requires segmentation and pitch determination. Segmentation is the process of determining the beginning and the end of a note. Prior art systems for segmenting notes in recordings of music rely on various techniques to determine note beginning points and end points. Techniques for segmenting notes include energy-based segmentation methods as disclosed in L. Rabiner and R. Schafer, "Digital Processing of Speech Signal," Prentice Hall: 1978, pp. 120-135 and L. Rabiner and B.H. Juang, "Fundamentals of Speech Recognition," Prentice Hall: New Jersey, 1993, pp. 143-149; voicing probability-based segmentation methods as disclosed in L. Rabiner and R. Schafer, "Digital Processing of Speech Signal," Prentice Hall: 1978, pp. 135-139, 156, 372-373, and T.F. Quatieri, "Discrete-Time Speech Signal Processing: Principles and Practice," Prentice Hall: New Jersey, 2002, pp. 516-519; and statistical methods based on stationarity measures or Hidden Markov models as disclosed in C. Raphael, "Automatic Segmentation of Acoustic Musical Signals Using Hidden Markov Models,"IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, No. 4, 1999, pp. 360-370. Once the note beginning and endpoints have been determined, the pitch of that note over the entire duration of the note must be determined. A variety of techniques for estimating the pitch of an audio signal are available, including autocorrelation techniques, cepstral techniques, wavelet techniques, and statistical techniques as disclosed in L. Rabiner and R. Schafer, "Digital Processing of Speech Signal," Prentice Hall: 1978, pp. 135-141, 150-161, 372-378; T.F.Quatieri, "Discrete-time Speech Signal Processing," Prentice Hall, New Jersey, 2002, pp. 504-516, and C. Raphael, "Automatic Segmentation of Acoustic Musical Signals Using Hidden Markov Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 4, 1999, pp. 360-370. Using any of

these techniques, the pitch can be measured at several times throughout the duration of a note. This resulting sequence of pitch estimates may then be used to assign a single pitch frequency to a note, as pitch estimates vary considerably over the duration of a note. This is true of ~~must~~ most acoustic instruments and especially the human voice, which is characterized by multiple harmonics, vibrato, aspiration, and other qualities which make the assignment of a single pitch quite difficult.

Please replace paragraph 32 beginning at page 10, with the following amended paragraph:

Fig. 1 is a block diagram of a system 10 suitable for accepting an input of a monophonicaudio signal. In a first alternative embodiment of the invention, the monophonic audio signal is a vocalized song. The system 10 provides an output of information for programming a corresponding ring tone for mobile telephones according to principles of the present invention. The system 10 has a telephony (or mobile) call handler 50, a ring tone sequence application 40 that transforms vocal input in accordance with the present invention, and a SMS handler 30. Input signal 5 from a source 2 is received at the call handler 50 for voice capture. The input signal would be of limited duration, for example, typically lasting between 5 and 60 seconds. Signals of shorter or longer duration are possible. The voice signal is then digitized and is then transmitted to the ring tone sequence subsystem 40. While the input shown here is an analog receiver such as an analog telephone, the input could also be received from ~~a~~ an analog-to-digital signal transducer. Further, instead of receiving an input signal over a telephone network, the input signal could instead be received at a kiosk or over the Internet.

Please replace paragraph 43 beginning at page 16, with the following amended paragraph:

The voicing probability measure 224 is defined as the point between the voiced and unvoiced portion of the frequency spectrum for one frame of the signal. A voiced signal is defined as a signal that contains only harmonically related spectral components whereas an unvoiced signal does not contain harmonically related spectral components and can be modeled as filtered noise. In the preferred embodiment, if $v = 1$ the frame of the signal is purely voiced; if $v = 0$, the frame of the signal is purely unvoiced.

Please replace paragraph 44 beginning at page 16, with the following amended paragraph:

The secondary feature estimation module 130, shown in Figure 2A, produces a set of time varying secondary features 135 ~~for~~ based on each of the features 125. Fig. 2C depicts a "secondary data structure" 135A used to store the secondary features 135 for one frame of the digitized input signal 15. The secondary feature estimation module 135 generates secondary features by taking short-term averages of the primary features 125 output from the primary feature estimation module 120. Short-term averages are typically taken over 2-10 frames. In a preferred embodiment, short- term averages are computed over three consecutive frames. Secondary features generated for each frame and stored in the secondary data structure 135A are:

Please replace paragraph 65 beginning at page 27, with the following amended paragraph:

The audio signal interface 406 includes a microphone 412, low pass filter 414 and analog to digital converter (ADC) 416 for receiving and preprocessing analog input signals. It also includes a speaker driver ~~418~~ 420 (which includes a digital to analog signal converter and signal shaping circuitry commonly found in "computer sound boards") and an audio speaker ~~420~~ 418.

Please replace paragraph 66 beginning at page 27, with the following amended paragraph:

The memory 410 stores an operating system 430, application programs ~~50~~, and the previously described signal processing modules. The other modules stored in the memory 410 have already been described above and are labeled with the same reference numbers as in the other figures.